Annals of the Academy of Romanian Scientists
Series on Science and Technology of Information
ISSN 2066-8526 Volume **8**, Number **1**/**2015**

COMPUTER-BASED POLYPHONIC ANALYSIS OF NOVELS

Cristian DASCĂLU[1], Ştefan TRAUŞAN-MATU[2]

Abstract. *This paper presents the theoretical foundations and some implementation details of a software product to assist the analysis of narrative in novels. The approach is based on the polyphonic model introduced by Mikhail Bakhtin for novels and extended by Tannen and Trăuşan-Matu to model collaboration and inter-animation in conversations. On this basis we have developed a semi-automated analysis system based on natural language processing techniques. An exemplification is presented for the case of the novel "Master and Margarita" by Mikhail Bulgakov.*

Keywords: Polyphonic model, inter-animation, Natural Language Processing, discourse analysis, narrative analysis

1. Introduction

This paper presents the theoretical foundations and some implementation details of a software application to assist the analysis of literary narratives. It is based on the polyphonic perspective introduced by Mikhail Bakhtin for novels and its extentions of Deborah Tannen [1] and Ştefan Trăuşan-Matu [2, 3] to model collaboration and inter-animation in conversations. It is considered that many "voices" are present in any verbal or textual communication, including both the voices of the effective participants, but also implicit ”voices”, for example, the ideas discussed in the text. Based on this model and the associated analytical method [4, 5] an analysis system has been developed based on natural language processing techniques. It allows a user to identify the "voices" (in the extended sense) of narratives and the interactions between them. The system is presented and exemplified in the case of the novel "Master and Margarita" by Mikhail Bulgakov [6]. The paper continues with a presentation of the polyphonic model of discourse. In the third section the system for analyzing novels is presented. The fourth section contains the conclusions.

2. The Polyphonic Model of Discourse

By analyzing the works of the Russian writer Fyodor Dostoyevsky, Mikhail Bakhtin emphasized a new discourse approach in novels, which may be considered as an important cause of the remarkable distinctive value of the

[1]MSc, Faculty of Automatic Control and Computers, University ”Politehnica” of Bucharest.
[2]Prof., PhD, Faculty of Automatic Control and Computers, University ”Politehnica” of Bucharest, Senior Researcher, Research Institute for Artificial Intelligence of the Romanian Academy, full member of the Academy of Romanian Scientists (stefan.trausan@cs.pub.ro).

celebrated Russian novelist. Bakhtin writes that one main feature of Dostoyevsky's novels is their polyphonic weaving of discourse, in other words, the weaving of multiple perspectives ("voices", if we follow the polyphonic music analogy), each being a natural expression of the character's well defined and autonomous conscience [7]. Thus, the novel no longer expresses the author's own beliefs, being rather the result of the interplay of the characters' ideas, beliefs and emotions. This usage of a plurality of voices, each retaining its individuality all through the novel is very similar, in Bakhtin's perspective, to the counterpoint technique utilized in polyphonic musical compositions [3].

An associated feature to the polyphonic dimension is that a specific trait of Dostoyevsky's novels is the existence of a real dialogue to which the author and the characters contribute with equal rights. There is no such thing as an all-knowing narrator - the puppeteer that is in control of the characters' destiny. In fact, the polyphonic effect is achieved by means of passing the narrator role between different characters. And, furthermore, there is no single absolute truth to be revealed at the end of the story – each participant has his own private, subjective, truth.

Coming back to the music analogy and according to Bakhtin's ideas, the polyphonic model was extended to any kind of discourse [8]. Every discourse can be seen as a collaborative activity that involves more than one voice, that is, two or more simultaneous "lines" of independent "melody" that, despite keeping an overall sense of unity, can temporarily be in disharmony [3, 8].

The central concept in the polyphonic model is inter-animation. It defines the way different voices combine on the longitudinal and vertical axis [3]. When the voices collaborate or negotiate a certain situation's resolution they are said to be in longitudinal inter-animation. This effect is most often achieved by means of repetition or elaboration. When the voices compete, dispute or reinforce a certain idea they are said to be in a state of vertical inter-animation which can tend either to unity, if the individual points of view are convergent, or difference, otherwise [3].

From the computational linguistics point of view, polyphony is a model for describing the structure of discourse. Rather than breaking down text to the sentence level, the polyphonic approach works with utterances and voices. An utterance can be anything from a single word to a whole paragraph or even a whole novel, as long as it conveys some sort of meaning and can be attributed to a voice. A voice can represent a certain character or even an individual idea and it is ultimately the central concept of the model [2-5, 8]. Throughout the discourse, voices come and go, interact, give birth to new ones, evolve separately although constantly harmonizing with each other – they can be viewed as the driving force of the narratives.

The polyphonic model has been applied successfully in the field of *Computer Supported Collaborative Learning* (CSCL) as a means of analyzing and assessing the activity in a collaborative chat-room - a medium for group learning or problem solving. In this scenario, peers debate, negotiate, dispute and elaborate a given topic having the ultimate goal of reaching a shared result. There are certain standard phases throughout this process: socialization (setting the common ground), problem understanding, exploration (peers discuss and negotiate different approaches) and, finally, result presentation [2-5].

Negotiation can work on two levels: one regarding the problem solving process and another regarding a common understanding of terms and their meanings. The latter precedes the former because in order for collaboration to arise all peers must have the same understanding of the terms that are going to be used. This is done explicitly, whenever confusion in semantics is noticed. On the other hand, this phenomenon might occur also implicitly: people usually have no problem in detecting the shift from the formal register to the satirical one, for example.

Inter-animation events can be used as a metric for assessing the quality of a collaborative learning session. Using this model we can detect if an utterance is on-topic, if it conforms to the general flow of discourse, if it is an elaboration of a previously stated idea or if it is a proposal of a totally new one, being either in a relationship of unity or one of difference with the rest [4, 9].

By tracking these events a graphical representation of the collaborative session can be generated which can be used to easily identify key moments in the discussion. This is of particular interest to the tutors overseeing the students' progress because it offers support for quick evaluation of their results [4, 9].

Although posing its difficulties, such as the need for spellchecking, co-reference resolution or adjacency pairs identification, the field of collaborative learning imposes a number of rules to the discourse structure that make its analysis a bit easier. First of all, the goal of such a discourse is very practical so convergence is usually sought after. Secondly, many of the topics that are going to be discussed are known beforehand. This means that analyzing chat logs is similar to conducting an experiment in a controlled environment: for example, in a debate-based collaboration process participants are each assigned a list of topics to defend and the measurement of their contribution is proportional to the number of utterances that involve said topics.

The polyphonic perspective was first introduced, by Bakhtin, having in mind literary texts, so incorporating it in a computational linguistics approach for analyzing such works is not far-fetched. Such an approach can be built upon the existing body of research in the field but as a first step some of the fundamental differences between the nature of literary text and collaborative learning session

transcripts must be noted. First of all, literary texts have an aesthetic function, are works of creativity and very often convey moral values. Their discourse does not necessarily converge to an easily definable point and, as a consequence, it is harder to track. Secondly, language is used in different ways and on different levels: it might be used to simply describe or to convey a broader view, referencing social or cultural phenomena – contextual information that is sometimes vital for the understanding of the text. Words can be used both with their literal or figurative meanings and may be used as hints to a greater picture.

3. Computer Supported Literary Text Analysis

The fact that literary texts use human language at its fullest potential, making use of complex rhetorical structures and a multitude of stylistic devices like metaphors, personification or allegory – to name only a few – makes a fully automated linguistic analysis practically (at least for the next years, in some opinions) impossible. In other words, the realistic goal, at the moment, is to design a tool to only assist and facilitate such an analysis. Using the polyphonic model as a framework for the text analysis, we can set the goal for our text analysis tool to be the discovery of central themes in a literary text (for example, a novel) and their interplay. This proves to be feasible because of the strong relationship between the concepts of voice, idea and theme. Still, user assistance will be needed in defining the literary themes but this step can also be semi-automated – we can guide our analysis by inspecting the most prominent semantic fields.

3.1 The Proposed Approach

In our approach, from the implementation point of view, the polyphonic discourse model was applied in the following manner:

- Each *paragraph* is treated as an *utterance*, its *emitter* being either the narrator or one of the characters
- A *voice* is a set of *related terms* – words that will most likely convey a certain meaning or theme.

In order to structure the input text in this way, a list of the characters that take part in the literary discourse must be provided. This is a task that can be semi-automated by using a Named Entity Recognizer to *gather a list of possible persons' names and then letting the user define the characters by providing each a name and a list of aliases*.

The next step is *utterance detection and emitter assignment*. As stated before, we start from the assumption that, in general, a paragraph change either signals the shift form one topic to another or is used to emphasize a certain phrase. Furthermore, during a dialogue, each character's reply is most of the time

contained in its own paragraph and there are ways to detect that the current paragraph is part of a dialogue – for example, it either starts with quotes or it has a special indentation. Using these assumptions we can assist the user in the utterance – emitter assignment task by parsing each paragraph and identifying cue phrases used by the narrator like "spoke", "added", "continued" and so on, using a predefined list. For example, let us examine the following paragraph from the novel The Master and Margarita by Mihail Bulgakov, Chapter 2 – "Pontius Pilate" (depicting the trial of Jesus of Nazareth)

> *" 'People of all kinds are streaming into the city for the feast-day. Among them there are magicians, astrologers, seers and murderers,' said the Procurator in a monotone. ' There are also liars. You, for instance, are a liar.' "* [6]

We have prior knowledge that the noun "procurator" refers to the character Pontius Pilate. We detect that it is the object of the cue-verb "spoke" so, consequently, we can make the assumption that this utterance belongs to Pilate. If the user sees this as unfit, he can manually edit this utterance – emitter assignment.

The last step is *voice definition* (we recall that, as mentioned above, we consider "voices" in a generalized way). For this we compute the frequencies of all the words that appear in the text and provide the user with a list sorted in descending order. This list can queried for synonyms using the WordNet dictionary's synset relationships. The user defines a voice by choosing a name and a set of associated terms. Using this information, individual occurrences are identified and a graph is plotted, showing utterances on the horizontal axis and emitters on the vertical axis. Using color codes, each voice is tracked through the discourse providing a visual representation for the inter-animation.

3.2 Case Study – Mihail Bulgakov's *Master and Margarita*

We are going to present the results of an analysis made using our polyphonic analysis tool on the first chapter of the novel *Master and Margarita* by Russian writer Mihail Bulgakov, which is considered as one of the most important writings of the 20th century and a accute satire of the staliniste regime [6]. Written between 1928 and 1940, unpublished until 1967, it tackles themes such as good and evil, love, cowardice and courage, atheism, and the moral state of the soviet society. Its narrative develops around two main themes: Woland's (the devil's personification) visit to Moscow and the story of Pontius Pilate and Yeshua (Jesus of Nazareth).

Chapter one introduces the theme of atheism. It depicts the encounter of two friends, Mikhail Alexandrovich, alias Berlioz, editor of a local literary journal, and Ivan Nikolayevich Ponyrev, alias Homeless, a young poet, with a peculiar foreigner who introduces himself as a professor specialized in black magic.

The three of them get into a discussion about the existence of God, Berlioz and Ivan denying his existence while the professor is trying to prove them wrong.

3.3 Performing the analysis

When using the text analysis tool, the first step of the implemented application is loading the input text. We can either load a previously saved project (which is an XML file containing the meta-data generated as a result of the analysis) or a brand new plain text file. The supported encoding is UTF-8.

The program's user interface is depicted in Figure 1. Notice the yellow bar on the left. That area will be used for displaying the name of the emitter of the utterance starting at that point. The middle section displays the text and the right panel contains the controls for defining meta-data like characters and voices.

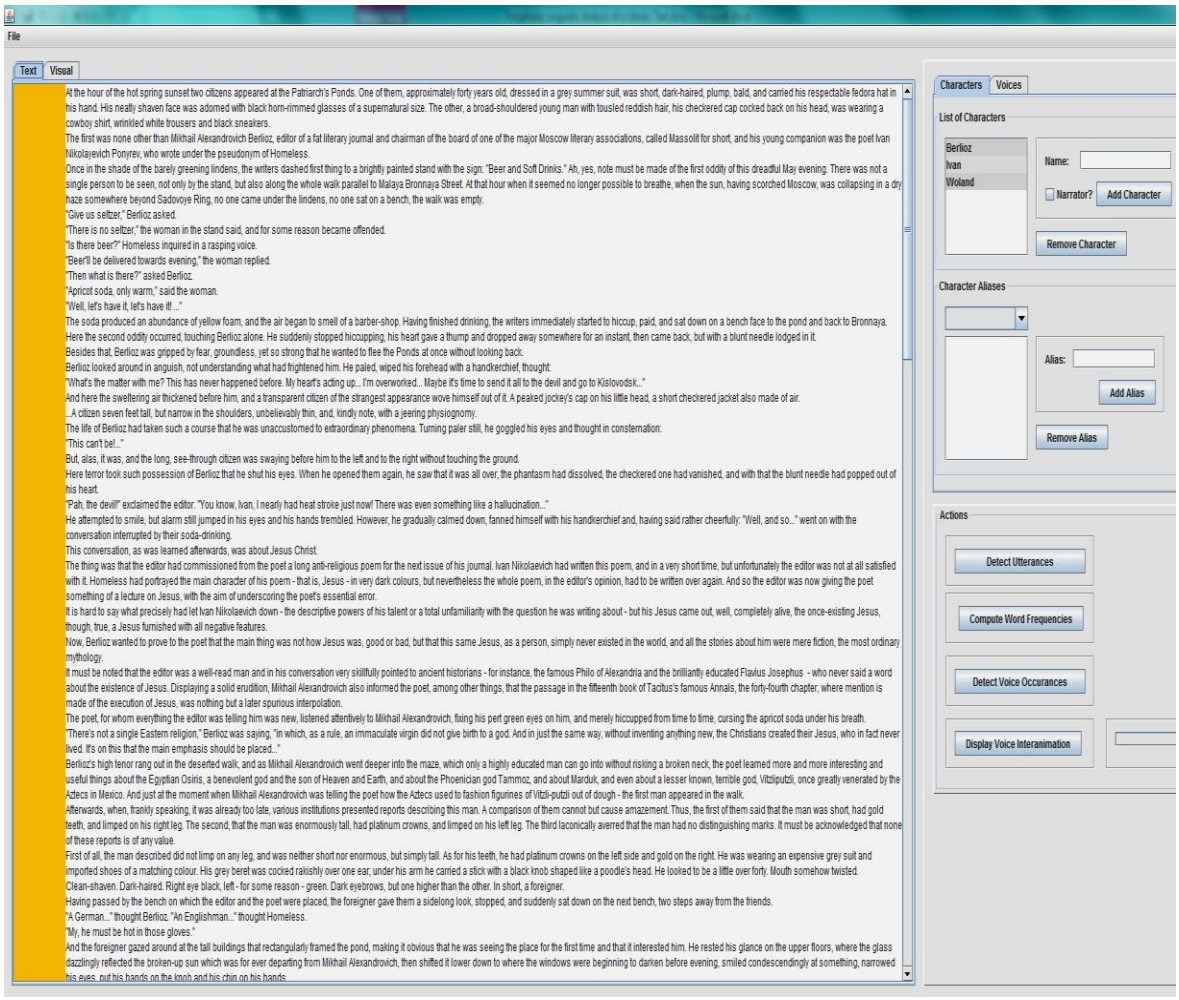

Fig. 1. The program's user interface.

The next step is the definition of the list of characters. This is done by using the controls from the right-side panel, more specifically in the "Characters" tab (see Figure 1).

The procedure for adding new characters is the following:

1. First we add the character by specifying his name – a name that from this point on will be used to identify it – this name will appear in the yellow bar on the left. It may not be its actual name because the character will later be detected using the list of aliases. Note the "Narrator?" checkbox. This is used to specify that that character is the narrator – which will be assigned for all the utterances that are not part of a dialogue.
2. For all characters excepting the narrator we must define the list of aliases by which they are referred throughout the text. This is done using the controls found in lower part of the panel depicted in Figure 1. Notice that for the character Ivan the name "Ivan" is also present in the aliases list. This is because only the aliases are taken into account when detecting utterance emitters.

After this step we can start the "Detect Utterances" procedure by clicking the corresponding button in the "Actions" panel from the lower right corner (see Figure 2). The result is depicted in Figure 3. Notice that the text is now displayed as a list of utterances with the emitters' names shown on the yellow bar.

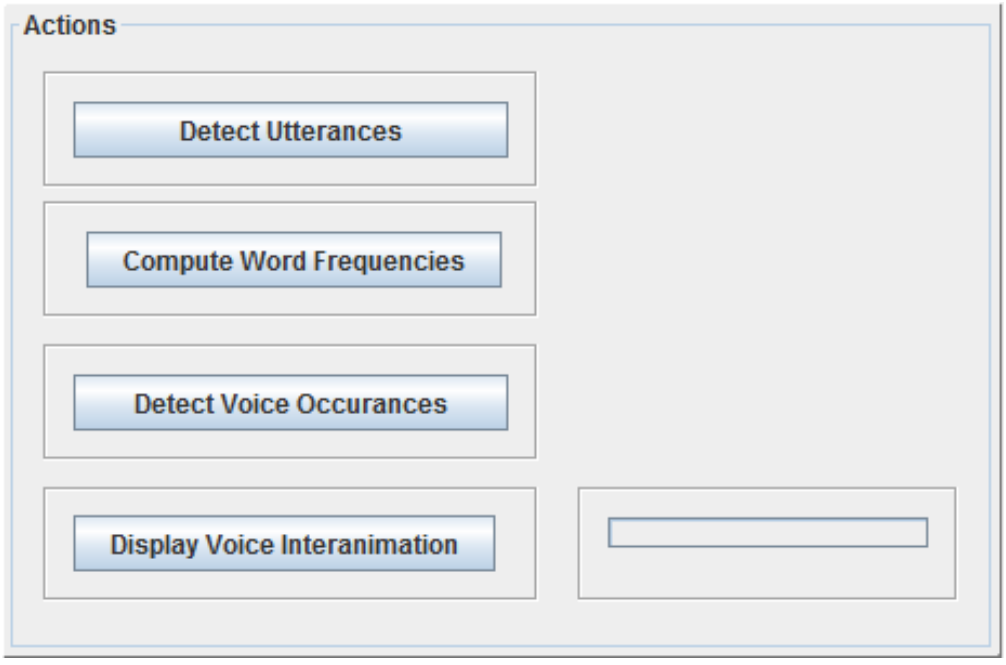

Fig. 2. The actions panel.

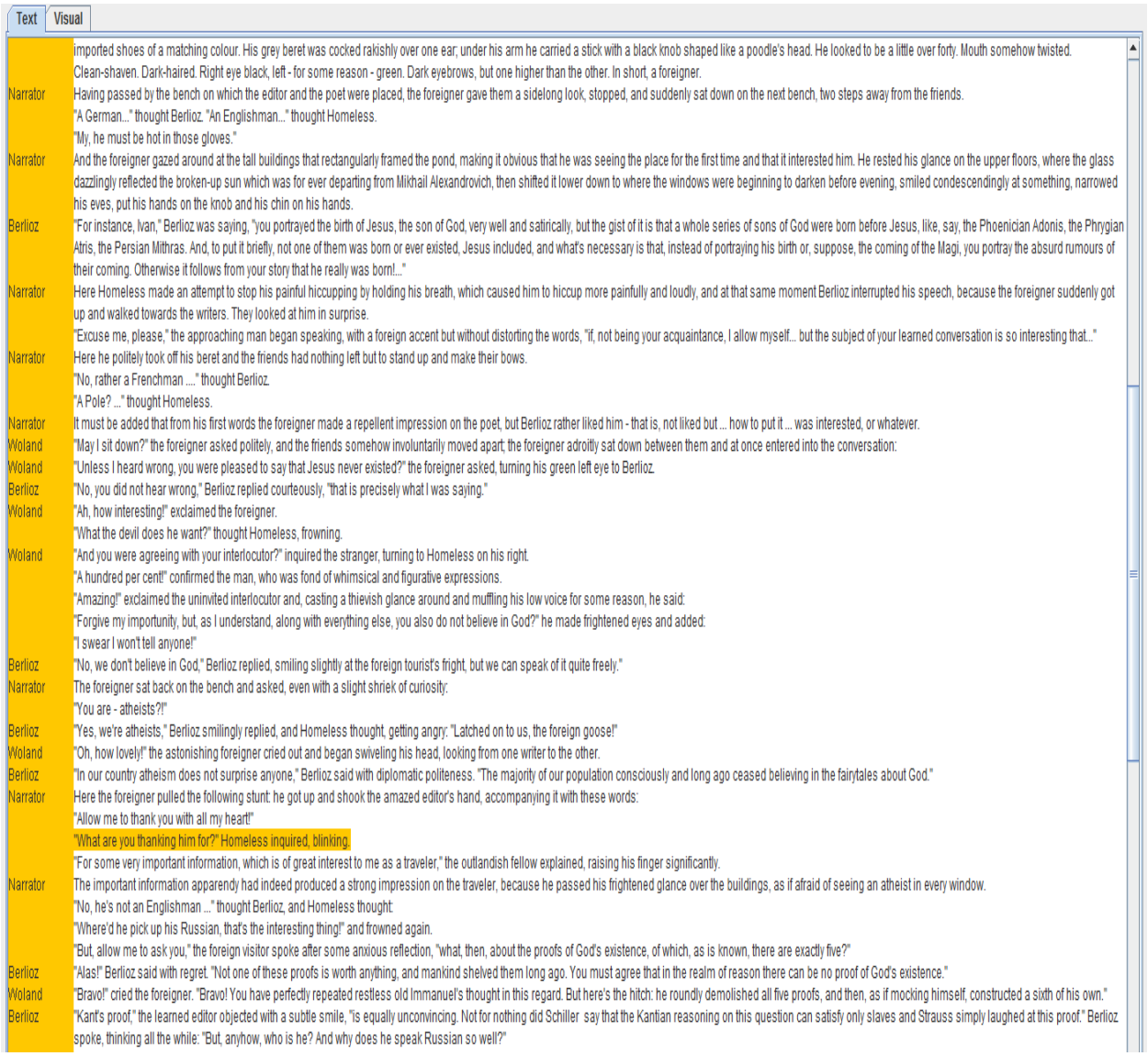

Fig. 3. The text divided at utterance level, annotated with emitter names.

By clicking inside the text display area, the whole utterance is highlighted in yellow, as a visual hint to help see the corresponding label from the bar on the left.

The next step is generating the list of words contained in the document. This is done by clicking the "Compute Word Frequencies" (see Figure 2). The result can be seen in the "Voices" tab of the upper right panel (see Figure 4). The list displays all the words found in the text in descending order of their frequencies and can be queried for synonyms.

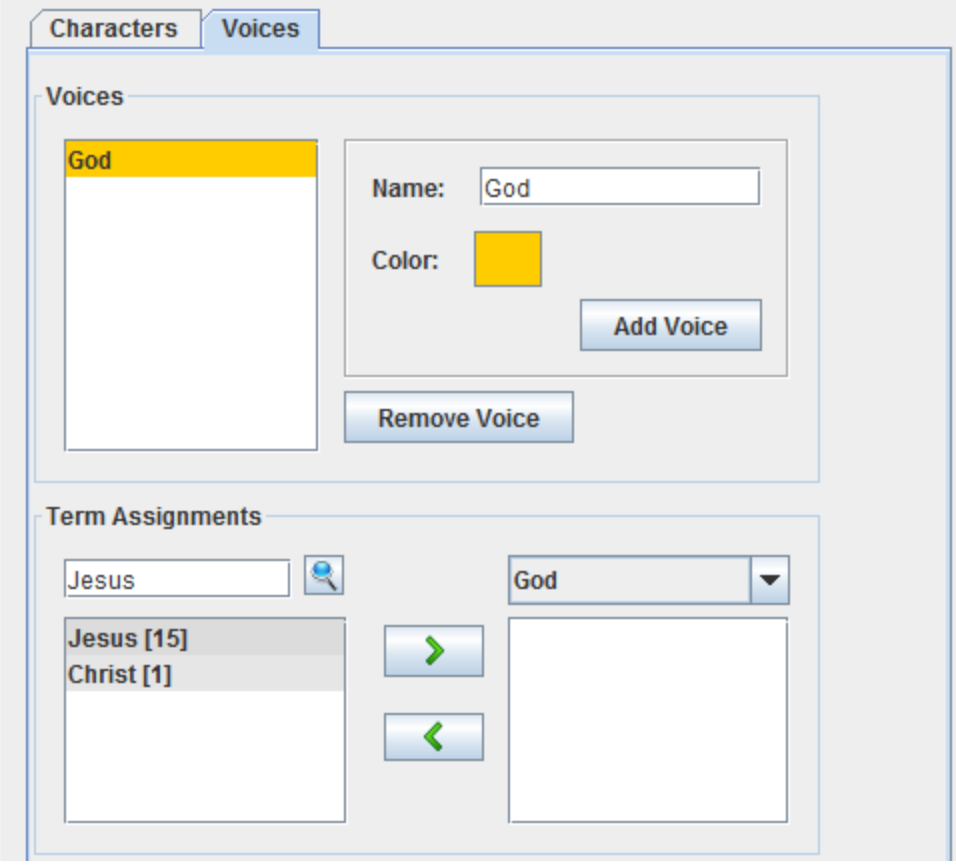

Fig. 4. The voices tab.

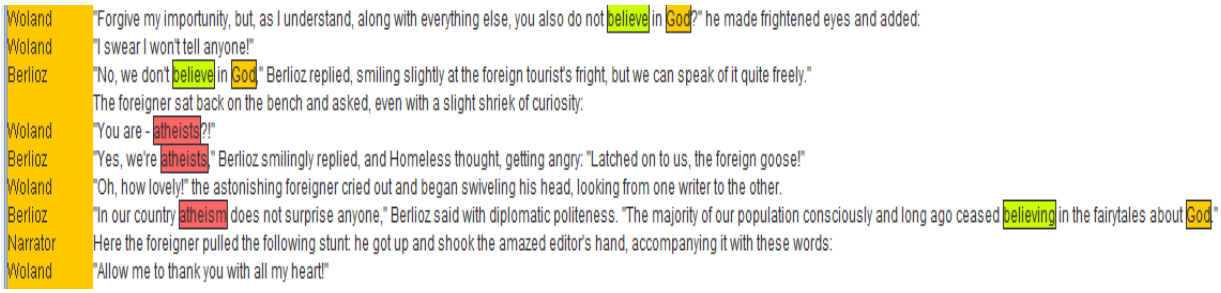

Fig. 5. Highlighted voice occurrences.

After computing the list of words we proceed to defining the desired voices. This is done using the controls from the "Voices" tab. To define a voice we must first choose a name and a color. After that we must assign some terms to it. For example, as show in Figure 4, if we want to capture the concept of divinity, we define a voice named God, pick a color for it and then assign terms to it by querying the list of words, selecting the appropriate ones and shuttling them to the list on the right side.

After finishing defining the voices we proceed to detecting their occurrences in the text. This is done by clicking the "Detect Voice Occurrences" button. As a result, each occurrence of a term assigned to a voice will be highlighted in the text, as shown in Figure 5.

So far, by inspecting the list of most frequent words we came across terms like *god*, *belief*, *atheism* and *proof*. We divided these into three voices: *God*, *Belief* and *Atheism*. The last step is to generate the inter-animation graph. This is done by clicking the "Display Voice Interanimation" button. For our text and the selected meta-data we obtain the inter-animation graph shown in Figure 6.

Figure 6 shows two very important visual cues:

- The beginning of the text has a fairly uniform structure – the theme of god's existence is disputed between Berlioz and Ivan.
- A change in structure occurs when the voice representing "atheism" first appears.
- A high level of inter-animation is detectable denoting a moment of tension in the dialogue.

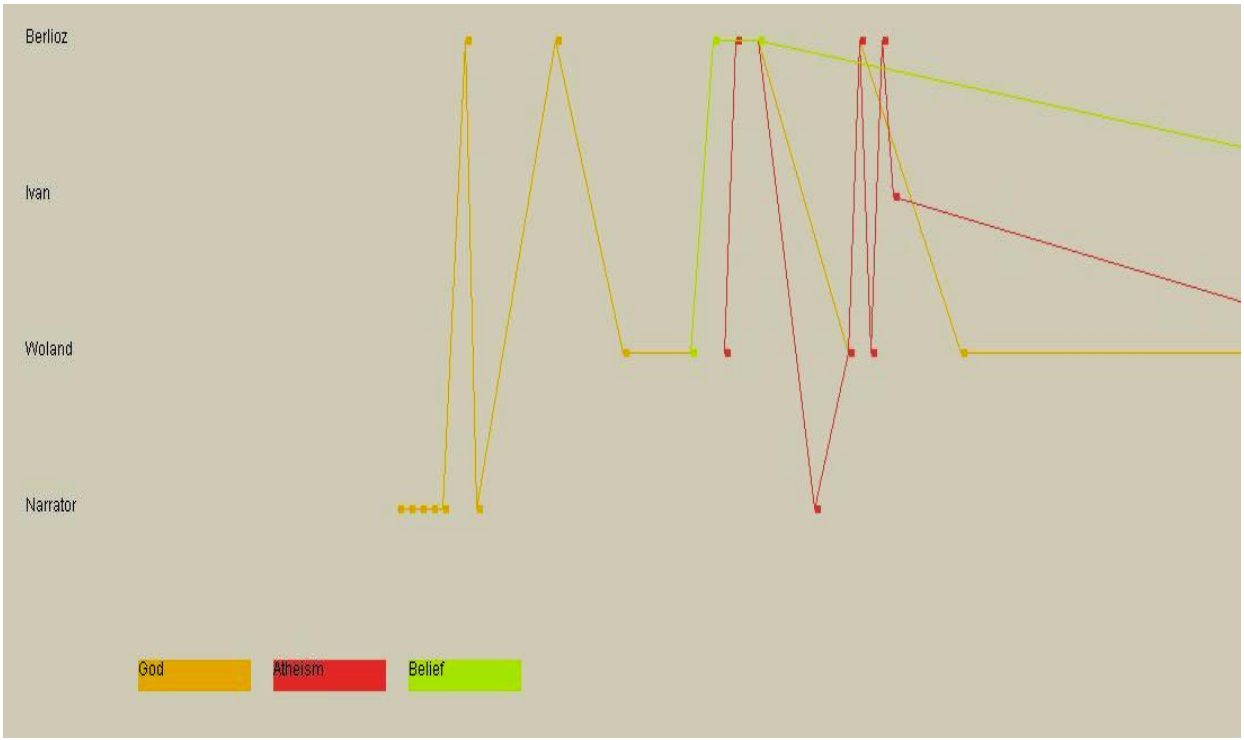

Fig. 6. The inter-animation graph.

The visual cue for the high density of voices' occurrences in a compact portion of text helps us discover the voices' inter-animation:

- Vertical
    - Unity: Berlioz and Ivan both admit that they are atheists
    - Difference: Berlioz and Ivan contradict Woland on the issue of God's existence.
- Longitudinal
    - Ivan backs up Berlioz' views about God's existence.

4. Conclusions

The polyphonic linguistic analysis tool can assist its user in structuring the text and providing it with a visual representation for its dynamics. It can draw the user's attention towards important aspects of the text using the degree voice inter-animation as a metric.

The actual computational linguistics technology state of the art most certainly cannot substitute a human expert. The presented system is designed to be used by a person who has some knowledge about the input text. It cannot tackle linguistic phenomena like polysemy or other stylistic devices like metaphors or speech register shifts.

R E F E R E N C E S

[1] Tannen, D.. (**1989**). *Talking Voices: Repetition, Dialogue, and Imagery in Conversational Discourse*, Cambridge University Press.

[2] Trausan-Matu, S. (**2010**a). The Polyphonic Model of Hybrid and Collaborative Learning, In Wang, F.,L., Fong., J., Kwan, R.C., *Handbook of Research on Hybrid Learning Models: Advanced Tools, Technologies, and Applications*, Information Science Publishing, Hershey, New York, pp 466-486.

[3] Trausan-Matu, S., Stahl, G., Sarmiento, J. (**2007**). Supporting Polyphonic Collaborative Learning, *E-service Journal*, vol. 6, nr. 1, Indiana University Press, pp. 58-74.

[4] Trausan-Matu, S. (**2010**). Automatic Support for the Analysis of Online Collaborative Learning Chat Conversations, in Philip M. Tsang, Simon K. S. Cheung, Victor S. K. Lee, Ronghuai Huang (Eds.): *Hybrid Learning, Third International Conference,* Beijing, LNCS 6248, Springer, pp. 383-394.

[5] Trausan-Matu, S (**2014**). Polyphonic Design, Conduct, Experience, and Evaluation in CSCL Chats, *Annals of the Academy of Romanian Scientists, Series on Science and Technology of Information*, Volume 7, Number 2/2014, pp. 21-34.

[6] Bulgakov, M. (**1967**). *The Master and Margarita*, Translated from the Russian by Michael Glenny. Published by Collins and Harvill Press, London.

[7] Bakhtin, M.M. (**1993**). *Problems of Dostoevsky's Poetics*, University of Minnesota Press, Minneapolis.

[8] Stefan Trausan-Matu, A Polyphonic Model, Analysis Method and Computer Support Tools for the Analysis of Socially-Built Discourse, *Romanian Journal of Information Science and Technology* 2-3/**2013**; 16:144-154.

[9] Trausan-Matu, S., Dascalu, M., Rebedea, T. (2014). PolyCAFe—automatic support for the polyphonic analysis of CSCL chats, *International Journal of Computer-Supported Collaborative Learning*, 06/**2014**; Volume 9(2), Springer, pp. 127-156.